Notes on function spaces, Hermitian operators, and Fourier series

S. G. Johnson, MIT Applied Mathematics

November 21, 2007

1   Introduction

In 18.06, we mainly worry about matrices and column vectors: finite-dimensional linear algebra. But into the syllabus pops an odd topic: Fourier series. What do these have to do with linear algebra? Where do their interesting properties, like orthogonality, come from?

In these notes, written to accompany 18.06 lectures in Fall 2007, we discuss these mysteries: Fourier series come from taking concepts like eigenvalues and eigenvectors and Hermitian matrices and applying them to *functions* instead of finite *column vectors*. In this way, we see that important properties like orthogonality of the Fourier series arises not by accident, but as a special case of a much more general fact, analogous to the fact that Hermitian matrices have orthogonal eigenvectors.

This material is important in at least two other ways. First, it shows you that the things you learn in 18.06 are not limited to matrices—they are tremendously more general than that. Second, in practice most large linear-algebra problems in science and engineering *come from* differential operators on functions, and the best way to analyze these problems in many cases is to apply the same linear-algebra concepts to the underlying function spaces.

2   Review: Finite-dimensional linear algebra

Most of 18.06 deals with finite-dimensional linear algebra. In particular, let's focus on the portion of the course having to do with *square* matrices and eigenproblems. There, we have:

- Vectors $\mathbf{x}$: column vectors in $\mathbb{R}^n$ (real) or $\mathbb{C}^n$ (complex).

- Dot products $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^H \mathbf{y}$. These have the key properties: $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2 > 0$ for $\mathbf{x} \neq \mathbf{0}$; $\mathbf{x} \cdot \mathbf{y} = \overline{\mathbf{y} \cdot \mathbf{x}}$; $\mathbf{x} \cdot (\alpha\mathbf{y} + \beta\mathbf{z}) = \alpha\mathbf{x} \cdot \mathbf{y} + \beta\mathbf{x} \cdot \mathbf{z}$.

- $n \times n$ matrices A. The key fact is that we can multiply A by a vector to get a new vector, and matrix-vector multiplication is *linear*: $A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y}$.

- Transposes A^T and adjoints $A^H = \overline{A^T}$. The key property here is that $\mathbf{x} \cdot (A\mathbf{y}) = (A^H\mathbf{x}) \cdot \mathbf{y}$ ... the whole reason that adjoints show up is to move matrices from one side to the other in dot products.

- Hermitian matrices $A = A^H$, for which $\mathbf{x} \cdot (A\mathbf{y}) = (A\mathbf{x}) \cdot \mathbf{y}$. Hermitian matrices have three key consequences for their eigenvalues/vectors: the eigenvalues λ are *real*; the eigenvectors are *orthogonal*; [1] and the matrix is *diagonalizable* (in fact, the eigenvectors can be chosen in the form of an *orthonormal basis*).

Now, we wish to carry over these concepts to functions instead of column vectors, and we will see that we arrive at Fourier series and many more remarkable things.

3   A vector space of functions

First, let us define a new vector space: the space of *functions* $f(x)$ defined on $x \in [0, 1]$, with the *boundary conditions* $f(0) = f(1) = 0$. For simplicity, we'll restrict ourselves to *real* $f(x)$. We've seen similar vector spaces a few times, in class and on problem sets.

This is clearly a vector space: if we add two such functions, or multiply by a constant, we get another such function (with the same boundary conditions).

Of course, this is not the only vector space of functions one might be interested in. One could look at functions on the whole real line, or two-dimensional functions $f(x, y)$, or even vector fields and crazier things. But this simple set of functions on $[0, 1]$ will be plenty for now!

4   A dot product of functions

To do really interesting stuff with this vector space, we will need to define the *dot product* (*inner product*) $f \cdot g$ of two functions $f(x)$ and $g(x)$.[2]

The dot product of two vectors is the sum of their components multiplied one by one and added (possibly with complex conjugation if they are complex). The corresponding thing for functions is to multiply $f(x)g(x)$ for each x and "add them up"—*integrate* it:

$$f \cdot g = \int_0^1 \overline{f(x)}g(x)dx. \tag{1}$$

For real functions, we can drop the complex conjugation of the $f(x)$. Equation (1) is easily seen to satisfy the key properties of dot products: $f \cdot g = \overline{g \cdot f}$; $f \cdot (\alpha g + \beta h) = \alpha f \cdot g + \beta f \cdot h$; $f \cdot f = \|f\|^2 > 0$ for $f \neq \mathbf{0}$.

---

[1]At least, the eigenvectors are orthogonal for distinct eigenvalues. In the case where one has multiple independent eigenvectors of the same eigenvalue λ, i.e. the null-space of $A - \lambda I$ is 2-or-more dimensional, we can always orthogonalize them via Gram-Schmidt as we saw in class. So, it is fairer to say that the eigenvectors can always be *chosen orthogonal*.

[2]The combination of a vector space and an inner product is called a *Hilbert space*. (Plus a technical condition called *completeness*, but that's only to deal with perverse functional analysts.)

Actually, the last property is the most tricky: the quantity

$$f \cdot f = \|f\|^2 = \int_0^1 |f(x)|^2 dx \qquad (2)$$

certainly *seems* like it must be positive for $f(x) \neq 0$. However, you can come up with annoying functions where this is not true. For example, consider the function $f(x)$ that $= 0$ everywhere except at $x = 0.5$, where $f(0.5) = 1$. This $f(x)$ is $\neq 0$, but its $\|f\|^2 = 0$ because the single point where it is nonzero has zero area. We can eliminate most of these annoyances by restricting ourselves to continuous functions, for example, although adding finite number of jump discontinuities is also okay. In general, though, this raises an important point: whenever you are dealing with functions instead of column vectors, it is easy to come up with crazy functions that behave badly (their integral doesn't converge or doesn't even exist, their $\|f\|^2$ integral is zero, etcetera). Functional analysts love to construct such pathological functions, and defining the precise minimal criteria to exclude them is quite tricky in general, so we won't try here. Let us colloquially follow the Google motto instead: *don't be evil*; in physical problems, pathological functions are rarely of interest. We will certainly exclude any functions where $\|f\|^2$ is not finite, or is zero for nonzero $f(x)$.

Equation (1) is not the only possible way to define a function dot product, of course. For example, $\int_0^1 f(x)g(x)x \, dx$ is also a perfectly good dot product that follows the same rules (and is important for problems in cylindrical coordinates). However, we will stick with the simple eq. (1) definition here, merely keeping in mind the fact that it was a *choice*, and the best choice may be problem-dependent.

5   Linear operators

A square matrix A corresponds to a linear operation $\mathbf{y} = A\mathbf{x}$ that, given a vector $\mathbf{x}$, produces a new vector $\mathbf{y}$ in the same space $\mathbb{C}^n$. The analogue of this, for functions, is some kind of operation $A f(x)$ that, given a function $f(x)$, produces a *new function* $g(x)$. Moreover, we require this to be a *linear* operation: we must have $A[\alpha f_1(x) + \beta f_2(x)] = \alpha A f_1(x) + \beta A f_2(x)$ for any constants α and β and functions f_1 and f_2.

This is best understood by example. Perhaps the simplest linear operation on functions is just the operation $Af(x) = af(x)$ that multiplies $f(x)$ by a real number a. This is clearly linear, and clearly produces another function! This is not a very interesting operation, however.

A more interesting type of linear operation is one that involves *derivatives*. For example, $Af(x) = df/dx = f'(x)$. This is clearly a linear operation (the derivative of a sum is the sum of the derivatives, etcetera). It produces a perfectly good new function $f'(x)$, as long as we don't worry about non-differentiable $f(x)$; again, *don't be evil*, and assume we have functions where A is defined. Another example would be $Bf(x) = A^2 f(x) = f''(x)$, the second-derivative operator. (Notice that A^2 just means we perform A, the derivative, twice.) Or we could add operators, for example $C = d^2/dx^2 + 3 \, d/dx + 4$ is another *linear differential operator*.

Of course, if we can make a linear operator out of derivatives, you might guess that we can make linear operators out of integrals too, and we certainly can! For example,

$Af(x) = \int_0^x f(x')dx'$. Notice that we put x in the integral limits (if we had used $\int_0^1$, then Af would have been just a number, not a function). This is also linear by the usual properties of integration. There are many other important integral operators, but it leads into the subject of *integral equations*, which is a bit unfamiliar to most students, and we won't delve into it here.

6 Adjoints of operators

The adjoint A^H of a matrix is just the complex conjugate of the transpose, and the transpose means that we swap rows and columns. But this definition doesn't make sense for linear operators on functions: what are the "rows and columns" of d/dx, for example? Instead, we recall that the *key property* of the adjoint (and the transpose, for real matrices) was how it interacts with dot products. In fact, handling matrices in dot products is essentially the whole reason for doing adjoints/transposes. So, we use this property as the *definition* of the adjoint: the adjoint A^H is the linear operator such that, for *all* $f(x)$ and $g(x)$ in the space,

$$f(x) \cdot [A\, g(x)] = [A^H f(x)] \cdot g(x). \tag{3}$$

That is, the adjoint is *whatever we must do to A in order to move it from one side to the other of the dot product*. (It is easily verified that, for ordinary matrices, this definition yields the ordinary conjugate-transpose A^H.)

This is best understood by an example. Take the derivative operator d/dx. We want to know what is its adjoint $(d/dx)^H$? From the perspective of matrices, this may seem like an odd question—the transpose of a derivative?? However, it is actually quite natural: for functions, our dot product (1) is an integral, and to move a derivative from one function to another inside an integral we *integrate by parts*:

$$f(x) \cdot \left[\frac{d}{dx}g(x)\right] = \int_0^1 f(x)g'(x)dx = f(x)g(x)|_0^1 - \int_0^1 f'(x)g(x)dx$$

$$= -\int_0^1 f'(x)g(x)dx = \left[-\frac{d}{dx}f(x)\right] \cdot g(x). \tag{4}$$

(We have dropped the complex conjugation, since we are dealing with real functions.) Note that, in going from the first line to the second line, we used the boundary conditions: $f(0) = f(1) = g(0) = g(1) = 0$, so the boundary term in integration by parts disappeared. If we compare eq. (3) with the first and last expressions of eq. (4) we find a wonderfully simple result:

$$\left(\frac{d}{dx}\right)^H = -\frac{d}{dx}. \tag{5}$$

That is, to move the derivative from one side to the other inside this dot product, we just *flip the sign* (due to integration by parts).

Before we go on, it is important to emphasize that eq. (5) is for *this* dot product and *this* function space. In general, the adjoint of an operator depends on all three things: the operator, the dot product, and the function space.

7   A Hermitian operator

Now that we have defined the adjoint A^H of an operator A, we can immediately define what we mean by a *Hermitian operator* on a function space: A is Hermitian if $A = A^H$, just as for matrices. Alternatively, based on the definition (3) of the adjoint, we can put it another way too: A is Hermitian if $f \cdot Ag = Af \cdot g$ for all f and g in the space.

What is an example of a Hermitian operator? From eq. (5), the *first* derivative certainly is *not* Hermitian; in fact, it is *anti*-Hermitian. Instead, let us consider the *second* derivative d^2/dx^2. To move this from one side to the other in the dot product (integral), we must integrate by parts *twice*:

$$
\begin{aligned}
f(x) \cdot \left[\frac{d^2}{dx^2} g(x) \right] = \int_0^1 f(x) g''(x) dx &= f(x) g'(x)|_0^1 - \int_0^1 f'(x) g'(x) dx \\
&= - \int_0^1 f'(x) g(x) dx = - f'(x) g(x)|_0^1 + \int_0^1 f''(x) g(x) dx \\
&= \int_0^1 f''(x) g(x) dx = \left[\frac{d^2}{dx^2} f(x) \right] \cdot g(x). \quad (6)
\end{aligned}
$$

Again, in going from the first line to the second, and in going from the second line to the third, we have used the boundary condition that both $f(x)$ and $g(x)$ are zero at the boundaries, so e.g. $f(x) g'(x)|_0^1 = 0$. Again, comparing the first expression to the last expression in eq. (6), we have found:

$$
\left(\frac{d^2}{dx^2} \right)^H = \frac{d^2}{dx^2}, \tag{7}
$$

i.e. that the *second* derivative operator *is* Hermitian!

There is another way to see the Hermitian property (7) of the second derivative by realizing that, from eq. (5):

$$
\frac{d^2}{dx^2} = \frac{d}{dx} \frac{d}{dx} = - \left(\frac{d}{dx} \right)^H \frac{d}{dx}. \tag{8}
$$

Whenever we have $A = -B^H B$ for any B, then A is Hermitian![3]

8   Eigenfunctions of a Hermitian operator

Now, let us consider the eigenfunctions $f(x)$ of the Hermitian second-derivative operator $A = d^2/dx^2$. These satisfy $Af(x) = \lambda f(x)$, i.e.

$$
f''(x) = \lambda f(x), \tag{9}
$$

---

[3]This derivation is perhaps a bit too glib, however, because when we take the derivative we aren't really in the same function space any more: $f'(x)$ does *not* satisfy $f'(0) = f'(1) = 0$ in general. A more advanced course would be more careful in making sure that doing the derivative twice has the same Hermitian property, because the boundary terms in the integration by parts may change.

along with the boundary condition $f(0) = f(1) = 0$.

What kind of function has a second derivative that is a constant multiple of itself? Exponentials e^{cx} would work, but they don't satisfy the boundary conditions—they are never zero! Or $\cos(cx)$ would work, but again it doesn't satisfy the boundary condition at $x = 0$. The only other alternative is $f(x) = \sin(cx)$, for which $f''(x) = -c^2 \sin(cx)$. The sine function works: $f(0) = 0$ automatically, and we can make $f(1) = 0$ by choosing c to be $n\pi$ for a positive integer $n = 1, 2, 3, \ldots$. Thus, the eigenfunctions and eigenvalues are (for an integer $n > 0$):

$$f_n(x) = \sin(n\pi x) \tag{10}$$

$$\lambda_n = -(n\pi)^2. \tag{11}$$

Notice that the eigenvalues are *real*, just as we would obtain for a Hermitian matrix. And, just as for a Hermitian matrix, we can show that the eigenfunctions are *orthogonal*: just doing a simple integral using a trigonometric identity $\sin a \sin b = [\cos(a - b) - \cos(a + b)]/2$, we obtain:

$$\int_0^1 \sin(n\pi x)\sin(m\pi x)dx$$

$$= \int_0^1 \frac{\cos[(n - m)\pi x] - \cos[(n + m)\pi x]}{2}dx$$

$$= \left(\frac{\sin[(n - m)\pi x]}{2(n - m)\pi} - \frac{\sin[(n + m)\pi x]}{2(n + m)\pi} \right)\bigg|_0^1 = 0 \quad (12)$$

for $n \neq m$.

This orthogonality relationship, and the fact that the eigenvalues are real, didn't fall out of the sky, however—we didn't actually need to do the integral to check that it would be true. It follows from *exactly* the same proof that we applied to the case of Hermitian matrices—the proof only used the interaction with the adjoint and the dot product, and never referred directly to the matrix entries or anything like that. As a review, let's just repeat the proofs here, except that we'll use $f_n(x)$ instead of $\mathbf{x}$. Suppose we have eigenfunctions $f_n(x)$ satisfying $Af_n = \lambda f_n$. To show that λ is real, we took the dot product of both sides of the eigenequation with f_n:

$$f_n \cdot Af_n = f_n \cdot (\lambda_n f_n) = \lambda_n(f_n \cdot f_n) = \lambda_n\|f_n\|^2$$
$$= (A^H f_n) \cdot f_n = (Af_n) \cdot f_n = (\lambda_n f_n) \cdot f_n = \overline{\lambda_n}\|f_n\|^2, \tag{13}$$

and thus $(\lambda - \bar{\lambda}_n)\|f_n\|^2 = 0$, which is only possible if $\lambda_n = \bar{\lambda}_n$: λ_n is real!

The key point here is that we *didn't need to do the integral* in eq. (12)—the eigenfunctions were *automatically orthogonal* due to the Hermitian property of d^2/dx^2. This is extremely useful, because most differential operators aren't so nice as the second derivative, and we usually don't know an analytic formula for the eigenfunctions (much less how to integrate them). Even for the sines, it is beautiful to see how the orthogonality is not just an "accident" of some trigonometric identities: it comes because sines are eigenfunctions of the second derivative, which is Hermitian.

9 The Fourier sine series

For Hermitian matrices, an important property is that a Hermitian matrix is always diagonalizable: the eigenvectors always form a basis of the vector space. As was first suggested by Joseph Fourier in 1822, the same holds true for the functions $\sin(n\pi x)$, which are eigenfunctions of the Hermitian operator d^2/dx^2. That is, "any" function $f(x)$ can be written as a sum of these eigenfunctions multiplied by some coefficients b_n:

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(n\pi x). \tag{14}$$

This is now known as a *Fourier sine series* for $f(x)$.

The precise meaning of "any" function and of the "=" in eq. (14) generated a century and a half of controversy, but when the dust had settled it turned out that Fourier was essentially right: the series in eq. (14) converges almost everywhere to $f(x)$ [except at isolated points of discontinuity, which we usually don't care about], as long as $\int |f(x)|^p$ exists (doesn't blow up, etc.) for some $p > 1$. The fascinating issue of the convergence of this sine series is discussed further with some numerical examples in another 18.06 handout on the web site.

Not worrying about convergence, let's consider the question of how to find the coefficients b_n. For matrices, if we have a basis of eigenfunctions forming the columns of a matrix S, we can write any vector $\mathbf{x}$ in this basis as $\mathbf{x} = S\mathbf{b}$, solving for the coefficients $\mathbf{b} = S^{-1}\mathbf{x}$. Solving for the coefficients is hard in general because it requires us to solve a linear equation, but in the special case of a Hermitian matrix then the eigenvectors can be chosen to be orthonormal, and hence $S = Q$ (Q unitary) and $\mathbf{b} = Q^{-1}\mathbf{x} = Q^H\mathbf{x}$. That is, each component b_m is just $b_m = \mathbf{q}_m \cdot \mathbf{x}$: we get the coefficients by taking dot products of $\mathbf{x}$ with the orthonormal eigenvectors. We can do *exactly* the same thing with the Fourier series, because the eigenvectors are again orthogonal.

That is, take the dot product of both sides of eq. (14) with $\sin(m\pi x)$:

$$\begin{aligned} f(x) \cdot \sin(m\pi x) &= \sum_{n=1}^{\infty} b_n \sin(n\pi x) \cdot \sin(m\pi x) \\ &= b_m \|\sin(m\pi x)\|^2 = b_m/2. \end{aligned} \tag{15}$$

To go from the first line to the second, we used the orthogonality relationship: $\sin(n\pi x) \cdot \sin(m\pi x)$ is *zero* unless $m = n$, so all but the mth term of the sum disappears. The final factor of $1/2$ is from the fact that $\int_0^1 \sin^2(m\pi x) = 1/2$. Thus, we have arrived at one of the remarkable formulas of mathematics:

$$b_n = 2 \int_0^1 f(x) \sin(n\pi x) dx, \tag{16}$$

which gives the coefficients of the Fourier sine series via a simple integral.

The key point is that this result is not limited to sine functions or to d^2/dx^2. It was a consequence of the fact that d^2/dx^2 is a Hermitian operator that allowed us to expand

any function in terms of the eigenfunctions (sines) and to get the coefficients via the orthogonality relationship. There are many, many other Hermitian operators that arise in a variety of problems, for which similar properties hold.[4]

10   The Fourier cosine series

The reason that we got sine functions was the boundary conditions $f(0) = f(1) = 0$ on our function space. To get cosine functions instead, we merely have to change the boundary conditions to zero *slope* instead of zero *value* at the boundaries: that is, we require $f'(0) = f'(1) = 0$.

Then, if we look for eigenfunctions of the second derivative, i.e. $f''(x) = \lambda f(x)$ with zero-slope boundaries, we are led to

$$f_n(x) = \cos(n\pi x) \tag{17}$$

$$\lambda_n = -(n\pi)^2. \tag{18}$$

Now, however, we must allow $n = 0$, since that gives a perfectly good non-zero eigenfunction $f_0(x) = 1$.

The eigenvalues are real, but are the eigenfunctions still orthogonal? We could check this by explicit integration as in eq. (12), but it is simpler and nicer just to check that d^2/dx^2 is still Hermitian. If we look at the integration by parts in eq. (6), it still works: the boundary terms look like $f(x)g'(x)$ and $f'(x)g(x)$, so they still are zero. So, we get orthogonality for free!

Again, we can expand "any" reasonable function in terms of the cosine eigenfunctions multiplied by some coefficients a_n, yielding the *Fourier cosine series*:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi x) \tag{19}$$

$$a_n = 2 \int_0^1 f(x) \cos(n\pi x) dx. \tag{20}$$

Again, we got eq. (20) by taking the dot product of both sides of eq. (19) with $\cos(m\pi x)$, which kills all of the terms in the series except for $n = m$, thanks to orthogonality. The a_0 term looks a bit funny: the extra factor of $1/2$ is there simply because $\| \cos(n\pi x) \|^2$ is $1/2$ for $n \neq 0$ but is 1 for $n = 0$. By including the $1/2$ in the series definition (19), we can use the same formula (20) for all the terms, including $n = 0$.

---

[4]Technically, the ability of the eigenfunctions to form a "basis" for the space, i.e. to be able to expand "any" function in terms of them, is unfortunately not automatic for Hermitian operators, unlike for Hermitian matrices where it is always true. A Hermitian operator on functions has to satisfy some additional properties for this property, the *spectral theorem*, to hold. However, the Hermitian operators that arise from physical problems almost always have these properties, so much so that many physicists and engineers aren't even aware of the counter-examples.

11   The Fourier series

The name *Fourier series* by itself is usually reserved for an expansion of $f(x)$ in terms of *both* sine and cosine. How do we get this? Zero boundaries gave sines, and zero-slope boundaries gave cosines; what condition does *both* sine and cosine satisfy? The answer is *periodic* boundaries. To keep the same formulas as above, it is convenient to look at functions $f(x)$ on $x \in [-1,1]$, with the dot product $f \cdot g = \int_{-1}^{1} f(x)g(x)dx$. We then require our functions $f(x)$ to be *periodic*: $f(-1) = f(1)$ and $f'(-1) = f'(1)$..

The periodic eigenfunctions are *both* $\sin(n\pi x)$ *and* $\cos(n\pi x)$ (i.e. two independent functions with the same eigenvalue $-n^2\pi^2$). We already know that $\sin$ is orthogonal to $\sin$ and $\cos$ to $\cos$, and $\sin$ is orthogonal to $\cos$ because the integral of an odd function multiplied by an even function is zero. Our operator d^2/dx^2 is still Hermitian: in the integration by parts [eq. (6)], we get boundary terms like $f(x)g'(x)|_{-1}^{1}$, which are zero because the periodicity means that $f(-1)g'(-1) = f(1)g'(1)$.

Finally, we can write out "any" function on $x \in [-1,1]$ as a sum of sines and cosines, and get the coefficients by orthogonality as above. However, we will mainly focus on the sine and cosine series, for simplicity.

12   A positive-definite operator

Let's go back to our original space of functions $f(x)$ for $x \in [0,1]$ with $f(0) = f(1) = 0$. Instead of d^2/dx^2, let's look at $-d^2/dx^2$. Since we just multiplied by -1, this doesn't change the Hermitian property or the eigenfunctions, and the eigenvalues just flip sign. That is, the eigenfunctions are still $\sin(n\pi x)$ for $n > 0$, and the eigenvalues are now $+(n\pi)^2$. Notice that the eigenvalues are all *positive*. In analogy with matrices, we can say that the operator $-d^2/dx^2$ must be *positive definite*.

However, it is unsatisfying to have to actually find the eigenvalues in order to check that the operator is positive definite. Solving for the eigenvalues may be hard, or even impossible without a computer, if we have a more complicated operator. We want a way to tell that the eigenvalues are positive just by looking at the operator, in the same way that we could tell that the eigenvalues were real just by integrating the operator by parts (to check that it is Hermitian). Fortunately, this is quite possible!

To check that an operator A is positive definite, we just need to check that $f \cdot Af > 0$ for all $f(x) \neq 0$. Typically, we can do this just by integrating by parts, and that works here as well:

$$f \cdot \left[-\frac{d^2}{dx^2} f \right] = -\int_0^1 f(x)f''(x)dx = -f(x)f'(x)|_0^1 + \int_0^1 [f'(x)]^2 \, dx \tag{21}$$
$$= \int_0^1 [f'(x)]^2 \, dx \geq 0.$$

So, we can see that the operator $-d^2/dx^2$ is at least positive semidefinite. For it to be positive definite, we need to make sure that eq. (21) is never zero for a non-zero $f(x)$. The only way for eq. (21) to equal zero is if $f'(x) = 0$, which means that $f(x)$ is a constant. But, with the boundary condition $f(0) = f(1) = 0$, the only constant $f(x)$

9

can be is zero. So, eq. (21) is always > 0 for $f(x) \neq 0$, and $-d^2/dx^2$ is positive-definite.

There is another way to see the same thing. From eq. (8), we saw that $A = -d^2/dx^2 = B^H B$, for $B = d/dx$. Any operator or matrix of the form $B^H B$ is automatically positive semidefinite (as we saw in class), and is positive-definite if B has full column rank (i.e., if B has no non-zero nullspace). Here, $Bf = f'(x) = 0$ only for $f(x) = 0$, due to the boundary conditions, by the same reasoning as above, so A is positive definite.

The ability to do this kind of analysis is tremendously important, because it allows us to say a lot about the eigenvalues of differential operators, even sometimes very complicated operators, without having to solve a horrible differential equation.

13 Fourier (sine) series applications

With diagonalizable matrices A, once we knew the eigenvalues λ and eigenvectors we could do all sorts of marvelous things: we could invert the matrix just by inverting the eigenvalues (if they are nonzero), we could take exponentials just by taking e^λ, we could compute powers A^n via λ^n (even square roots of matrices), and so on. The basic strategy was always the same: write an arbitrary vector in the basis of the eigenvectors, and then A acting on each eigenvector acts just like a simple number λ. This is the whole point of eigenvectors: they turn horrible things like matrices into simple numbers. Things work in exactly the same way with differential operators! Below are a couple of nice examples.

13.1 The diffusion equation

If we have a system of linear differential equations $d\mathbf{x}/dt = A\mathbf{x}$ with an initial condition $\mathbf{x}(0)$, we saw in class that the solution was just $\mathbf{x}(t) = e^{At}\mathbf{x}(0)$. The matrix exponential at first seemed rather strange, but in terms of the eigenvectors it is simple: we first write $\mathbf{x}(0)$ in the basis of eigenvectors, then multiply each eigenvector by $e^{\lambda t}$.

Now, let's look at the analogous problem of a *diffusion equation*:

$$\frac{\partial}{\partial t} f(x,t) = \frac{\partial^2}{\partial x^2} f(x,t), \tag{22}$$

where the initial condition $f(x,0)$ is given, and at each time t we have $f(0,t) = f(1,t) = 0$ (the same boundary conditions as for the sine series, hint hint). This equation is used to describe diffusion, e.g. f could be the concentration of salt in a solution of water; if you start out with a high concentration of salt in one region at $t = 0$, it should diffuse to other regions over time. Or, it could also describe heat conduction, where f might be a temperature difference.

In any case, eq. (22), if we squint at it, is in the same form as $d\mathbf{x}/dt = A\mathbf{x}$, with $\mathbf{x}$ replaced by the function f and the matrix A replaced by the operator $\partial^2/\partial x^2$. Thus, by analogy, we should be able to immediately write down the "solution"

$$f(x,t) = e^{t\frac{\partial}{\partial x^2}} f(x,0). \tag{23}$$

But wait a minute, what is the exponential of a second derivative? What does this even *mean*? We could definite it by a power series expansion, just as for matrices, but it is much more useful to think about what it does to eigenfunctions:

$$e^{t\frac{\partial}{\partial x^2}}\sin(n\pi x) = e^{-(n\pi)^2 t}\sin(n\pi x). \tag{24}$$

That is, we may not know what the exponential of a second derivative is in general, but we surely know what it does to an eigenfunction: For an eigenfunction $\sin(n\pi x)$, the second derivative $\partial^2/\partial x^2$ must act just like a *number*, the eigenvalue $-(n\pi)^2$.

Now that we know what the solution is for the eigenfunctions, however, we are done! We just expand the initial condition $f(x,0)$ via the sine series, multiply each $\sin(n\pi x)$ by $\exp(-(n\pi)^2 t)$, and voila, we have the solution:

$$f(x,t) = \sum_{n=1}^{\infty} b_n e^{-(n\pi)^2 t}\sin(n\pi x), \tag{25}$$

where $b_n = 2\int_0^1 f(x,0)\sin(n\pi x)$ as in eq. (16).

Of course, even if we can compute the b_n coefficients analytically [if $f(x,0)$ is simple enough to integrate], we probably can't add up the series (25) by hand. But it is no problem to add up 100 or 1000 terms of the series on a computer, to get the solution as a function of time.

Also, we have learned quite a bit just by writing down the series solution eq. (25). First, notice that all of the terms go exponentially to zero: $f(x,\infty) = 0$ (with these boundary conditions, everything "diffuses away" out of the boundaries). Second, notice that the large-n terms decay faster than the smaller-n terms: in diffusion problems, fast oscillations (large n) quickly smooth out, and eventually we are dominated by the $n = 1$ term $\sin(\pi x)$.

13.2   Poisson's equation

With matrices, solving $A\mathbf{x} = \mathbf{b}$ for $\mathbf{x}$ is fairly hard, it requires us to do elimination or something similar ($\sim n^3$ work). However, if A is diagonalizable and we know the eigenvectors, it is no problem: we just expand $\mathbf{b}$ in the eigenvectors, multiply each by $1/\lambda$, and we are done.[5]

Now, let's look at an analogous problem for a linear operator, *Poisson's equation*:

$$\frac{d^2}{dx^2}f(x) = g(x), \tag{26}$$

in which we are given $g(x)$ and want to find $f(x)$ such that $f(0) = f(1) = 0$. Again, this is "just" solving a linear equation, except that we have functions instead of vectors. Again, if we had eigenfunctions, we would know what to do: the solution to

$$\frac{d^2}{dx^2}f(x) = \sin(n\pi x), \tag{27}$$

---

[5]If A is not invertible, then we require $\mathbf{b}$ to be in the column space, which means that it must be in the span of the $\lambda \neq 0$ eigenvectors. And if we have a solution, it is not unique: we can add anything in the nullspace, which is the span of the $\lambda = 0$ eigenvectors. So, the eigenvectors still tell us everything even in the non-invertible case.

is clearly just $f(x) = \sin(n\pi x)/[-(n\pi)^2]$. That is, for eigenfunctions, we just take $1/\lambda$ as usual to solve the problem. And now we know what to do for an arbitrary $g(x)$: we expand it in a sine series, and divide each term by the eigenvalue. That is, the general solution is

$$f(x,t) = \sum_{n=1}^{\infty} \frac{b_n}{-(n\pi)} \sin(n\pi x), \tag{28}$$

where $b_n = 2\int_0^1 g(x)\sin(n\pi x)$ as in eq. (16).

In this case, the operator d^2/dx^2 is invertible because all the eigenvalues are non-zero. In the next problem set, you will think a little about the case of $f'(0) = f'(1) = 0$ boundary conditions, where one obtains a cosine series and there is a zero eigenvalue.

14 More examples of Hermitian operators

In class, we focused on the case of the operator d^2/dx^2, which has the virtue that its eigenfunctions can be found analytically, and so we can see properties like orthogonality and real eigenvalues explicitly. However, the real power of looking at differential operators in this way comes for problems that you *cannot* solve analytically (or where the solutions are very difficult). You can learn so much just by integrating by parts a couple of times, which is a lot easier than solving a partial differential equation. In this section, we'll give a few example of other operators that can be analyzed in this way.

In homework, you'll look at the operator $Af = -\frac{d}{dx}[w(x)df/dx]$ for some function $w(x) > 0$, which is also straightforwardly shown to be Hermitian positive-definite. This operator arises in many problems; for example, it appears when studying electromagnetic waves going through different materials, where f is the magnetic field and $1/w(x)$ is the square of the refractive index.

In quantum mechanics, one studies eigenproblems of the form (in one dimension)

$$\left[-\frac{d}{dx}^2 + V(x) \right] \psi(x) = E\psi(x), \tag{29}$$

where the eigenfunction ψ is a quantum probability amplitude, the eigenvalue E is the energy, and $V(x)$ is some potential energy function. This is clearly Hermitian, since it is the sum of two Hermitian operators: $-d^2/dx^2$, and $V(x)$ (which just multiplies the function by a real number at each point, and is trivially Hermitian). So, we immediately obtain the result that the energy E is real, which is good! (What would a complex energy mean?) And the eigenfunctions $\psi(x)$ are orthogonal, which turns out to have important physical consequences for the probabilities. And we learn all of this without solving it, which might be quite difficult depending on $V(x)$ is. One famous case is where $V(x) = x^2$, in which case the eigenfunctions are called *Hermite functions* and are important in many fields of science and mathematics.

Another important differential operator is:

$$Af(x) = \left[-\frac{d^2}{dx^2} - \frac{1}{x}\frac{d}{dx} + \alpha^2 \right] f(x) \tag{30}$$

for functions with $f(0) = f(1) = 0$ and some constant α. This operator A is Hermitian positive-definite under the inner product $f \cdot g = \int_0^1 f(x)g(x)x\,dx$, which arises in cylindrical coordinates (think $x = r$). The eigenfunctions are famous functions known as *Bessel functions*, but even if you don't know what these are you now know that they are orthogonal under this inner product, and that their eigenvalues are positive.

The preceding examples were for one-variable functions $f(x)$. Of course, there are lots of interesting problems in two and three spatial dimensions, too! We'll just give one example here. Suppose you have a drum (the musical instrument), and the drum head is some shape. Let $f(x, y)$ be the height of the drum head at each point (x, y), with $f(x, y) = 0$ at the boundaries of the drum head. One interesting problem is to find the *standing-wave* modes, which are solutions $f(x, y)\sin(\omega t)$ that oscillate at some fixed frequency ω. The standing-wave solutions satisfy the equation:

$$\left[-\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \right] f(x, y) = \omega^2 f(x, y). \tag{31}$$

Again, the linear operator on the left-hand side is Hermitian positive-definite with the inner product $f \cdot g = \iint f(x, y)g(x, y)dx\,dy$. The proof is almost exactly the same as for $-d^2/dx^2$ in one dimension: we just integrate by parts in x for the $\partial^2/\partial x^2$ term, and in y for the $\partial^2/\partial y^2$ term. This tells us that ω is real, which is good because it means that the solutions are oscillating instead of exponentially decaying or growing (as they would for complex ω). Solving for the eigenfunctions $f(x, y)$ explicitly, however, is quite hard in general (unless the drum head has a special shape, like square or circular). However, we do know that the eigenfunctions are orthogonal and form a basis for arbitrary functions $f(x, y)$. If you hit the drum, the frequencies that you hear are determined by taking the function that you hit the drum with (i.e., where you press down) and expanding it in the eigenfunctions. Hitting it at different points excites different eigenfunctions with different coefficients, and produces different tones.

MIT OpenCourseWare
http://ocw.mit.edu

18.303 Linear Partial Differential Equations: Analysis and Numerics
Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.